Deliberation and Identity Rules: The Effect of Anonymity, Pseudonyms, and Real-name Requirements on the Cognitive Complexity of Online News Comments

Alfred Moore, Politics, University of York

Rolf E. Fredheim, NATO Strategic Communications Centre of Excellence, Riga

Dominik Wyss, Politikwissenschaft, University of Stuttgart

Simon Beste, Social Sciences, Goethe University, Frankfurt

[Version accepted for publication by Political Studies on 6 Nov 2019.]

Introduction

Does anonymity — the ability to act publicly while concealing one's identity — promote or undermine public deliberation online? There are two general and opposed responses to this question. One is that anonymity is dangerous because it enables the evasion of accountability. Discussing a case of misogynistic abuse on an anonymous chat forum, Martha Nussbaum observes that 'the ability to create a new world in which [the abusers] exercise power and the women are humiliated depends on their ability to insulate their Internet selves from responsibility in the real world, while ensuring real-world consequences for the woman' (Nussbaum 2010, 85). In a similar vein, Levmore notes that online anonymity allows 'communication without retribution' (Levmore 1996: 2192-3). If the value of public deliberation is rooted in its insulation from sources of coercive power (Habermas 1996), then on this view anonymity would seem to undermine it by introducing an important power asymmetry into contexts of public communication. On the other hand, by insulating citizens from soft social pressures or hard sanctions and punishments, anonymity can enable people to speak in public with greater

sincerity. Timur Kuran thus suggests that anonymity can limit the dangers of 'preference falsification', that is, of misrepresenting one's preferences in order to conform to perceived public opinion. Preference falsification, on Kuran's view, can lead to the suppression of minority opinion, with the consequence that minority views either weaken and disappear, or break out in a sudden and radical bursts (Kuran 1997, 261). This claim is echoed by Danielle Allen in her discussion of Creon's regime in Sophocles' Antigone, in which it was only the anonymous chorus who were able to speak truth to Creon's power (though he did not listen), and from which she concludes that regimes of enforced public silence may look stable, but are prone to 'rapid, radical change' (Allen 2010, 117). So anonymity can on the one hand introduce power asymmetries and the strategic use of speech in the public sphere, and at the same time offer release from demands for social conformity that can themselves reflect power asymmetries (Asenbaum 2018). A reasonable, if not especially helpful, general answer to the question of the effect of anonymity on public deliberation would thus seem to be: it depends.[1] Assessing the question of the relationship between identity rules and the deliberative quality of the public sphere would thus seem to be something to be done on a case-by-case basis (and we present such an empirical analysis in this paper).

Before we go on, however, we should make an important conceptual point. It is common to frame the question of the deliberative value of anonymity in terms of a simple dichotomy, a trade-off between the goods and bads of identifiability and anonymity. This is how it is framed by the theorists mentioned above. This is also how it has appeared in debates on policy with respect to online discussion environments. For instance, Facebook's (then) marketing

[1] As Asenbaum (2018, 470) observes, the identity performances associated with anonymity are highly context-dependent.

director, Randi Zuckerberg, said during a panel discussion on social media in 2011 that 'anonymity on the internet has to go away. … People behave a lot better when they have their real names down,' she continued. 'I think people hide behind anonymity and they feel like they can say whatever they want behind closed doors' (quoted in Chun 2015: 105). Indeed, it was this same framing of the problem that led the Huffington Post a couple of years later to stop accepting anonymous comments in order to reduce trolling and verbal abuse, one of the changes whose effects are discussed in our study. This is also how it has been framed by a number of empirical studies looking at the effect of anonymity on online commenting (Manosevich et al 2014, 1180; Janssen and Kies 2005; Towne and Herbsleb 2012). Yet a number of theorists have recently highlighted the conceptual complexity of anonymity and its productive and communicative dimensions (Asenbaum 2018, Moore 2018, Veliz 2018). The case we examine in this paper gives us an opportunity to empirically address these claims.

In this paper we aim to take a step towards answering the broad question of the effects of identity rules on online deliberation by focusing more narrowly on the effects of changing modes of user identification on the cognitive complexity of comments on online news articles. A change in the rules of the Huffington Post (HuffPo) comment forum presented a natural opportunity to explore this question. Its comment forum was characterised by what we call non-durable pseudonymity (with a lot of trolling and large numbers of multiple or spoof accounts), until December 2013, when HuffPo changed the rules for its comment forums to require participants to authenticate their accounts through Facebook. This created an environment in which users did not have to use their real name with their comments, but nor could they so easily set up new accounts. In a second change, in June 2014, HuffPo adopted Facebook as the commenting system. This meant that HuffPo usernames were replaced with

user's Facebook names and avatars, approximating a real-name environment. This enabled a large-scale 'before and after' analysis. We collected over 45 million comments on more than 50,000 articles featured on the HuffPo front page between January 2013 and March 2015, and analysed them for cognitive complexity, which, we suggest below, is a good proxy for levels of argumentation and justification, which in turn are core elements of deliberative quality.[2]

We begin by reviewing the existing literature on the analysis of online news commenting and establishing the predominant expectations that would be associated with a shift from anonymous to real-name environments. We then describe the case of commenting on the Huffington post and the gathering of data, paying particular attention to the concept of anonymity, and differentiating between anonymous, pseudonymous and real-name environments. We then introduce our distinctive method for analysing one particular dimension of deliberative quality at scale through cognitive complexity. We then present our results, which both confirm and confound widespread expectations that real-name environments tend to better quality deliberation. We conclude with some reflections on the implications of these findings for online institutional design.

Analysing Comments Below the Line

Over the last decade commenting on online news publications has become a prominent part of the networked public sphere. Research into online commenting has ranged widely (see Reagle 2015; Stroud et al. 2017), though according to one recent survey, research has tended to focus on aspects of online institutional design and their effect on communicative behaviour

[2] We are not claiming that CC is a proxy for deliberative quality in the round. Our analysis of linguistic indicators of justification should be taken as a cue to further, more nuanced explorations using established qualitative approaches to discourse analysis.

within the forum (Friess and Eilders 2015). Within this general orientation to design choices and their relation to communication, identity rules have been a particular focus of attention. Many of these studies are framed in terms of the trade-off mentioned above, between the goods and bads of anonymity, and the goods and bads of real-name environments. Thus, a common finding has been that anonymity increases participation but lowers the quality of discussion (Rowe 2014; Towne and Herbsleb 2012). Conversely, online discussion spaces requiring users to authenticate their identities have lower rates of participation, but are associated with improvements in civility, rationality, and sincerity (Coleman and Moss 2012, 8; Janssen and Kies 2005, 231; Towne and Herbsleb 2012).

The most commonly investigated relation is between anonymity and civility, typically framed by the expectation that '[p]eople who are able to post anonymously (or pseudonymously) are far more likely to say awful things… Speaking from behind a blank wall that shields a person from responsibility encourages recklessness — it's far easier to simply hit the "send" button without a second thought under those circumstances' (Foxman and Wolf 2013, 114). Regarding civility, it has been argued on the basis of both psychological experiments and qualitative analysis of commenting forums that moving towards real-name environments is likely to improve civility. The dominant claim in the social psychology literature is that anonymity online promotes 'deindividuation' (see Reicher et al. 1995; Lea et al. 1994). The idea is that in certain social situations individuals lose connection to personal and social constraints, the so-called the 'online disinhibition effect' (Suler 2005). The loss of inhibition experienced with anonymity can be a good or bad thing, depending on the inhibition in question. Anonymity can release people from inhibitions linked to social and economic status (see Bowker and Tuffin 2003; Kim 2006). Anonymity can also enable people to test out arguments and change

their positions (Black 2011: 14). Yet others have emphasised the negative effects of such release from constraints. Sia et al. (2002) claim on the basis of a laboratory study that anonymity decreases 'social presence', which in turn increases polarization of discourse. Indeed, the predominant view is that anonymous environments are associated with uncivil, abusive or 'anti-normative' communication (Polat and Pratchett 2009; Lea et al 2001).

Studies focusing more broadly on deliberative quality lead to similar expectations. By closely examining a sample of comments from two regional US newspapers, Manosevitch and Walker (2009) were able to code comments and assess them for respectful engagement, exchange of arguments, accuracy of information, and general deliberative quality. Santana compared samples of comments from American newspapers that did and did not allow anonymous commenting, and found that 'Anonymous commenters were significantly more likely to register their opinion with an uncivil comment than non-anonymous commenters' (Santana 2014: 27). Rowe (2014) compared comments on the Washington Post's website (which allows anonymity) and discussions of the same topic on the Washington Post's Facebook discussion board (a real-name environment), and concludes that the website discussion is 'significantly more likely to be uncivil' (Rowe 2014: 121). Undertaking a similar comparison between comments on a news website and comments on the same story on a Facebook page, Hille and Bakker (2014) also find that by 'discouraging anonymous responses, the quality of comments improved but above all the quantity of comments decreased after outsourcing comments to Facebook' (Hille and Bakker 2014, 563). And based on an online discussion among Korean voters in the run up to an election, Rhee and Kim (2009) found that when users could post anonymously they were more engaged, in terms of quantity, but that when there were identity cues (which does not necessarily mean real name, which can in some con-

texts reveal a lot less than descriptions like ‘38 year old mother of two’) outside observers evaluated the quality of the discussion more highly.

Our study makes two distinctive contributions to this discussion. First, we explicitly thematise pseudonymous communication, in which commenters have stable identities that are not their real names and are not easily linked to other social contexts. Much of the existing work on identity rules and online commenting reproduces a simple dichotomy between anonymity and real-name identity (Hille and Bakker 2014; Janssen and Kies 2005; Rowe 2015; Santana 2014; Towne and Herbsleb 2015), and pays relatively little critical attention to those terms. Thus, for instance, Towne and Herbsleb use terms like ‘anonymous’ and ‘pseudonymous’ interchangeably (Towne and Herbsleb 2012, 108), as do Foxman and Wolf (2013, 114). Yet the fact that these researchers also invoke terms like ‘complete anonymity’ (Santana 2014, 28), and ‘fully anonymous’ (Towne and Herbsleb 2012, 108), indicates that there is recognition of the need for a further distinction. We will thus develop below an explicit theoretical account of identity rules, drawing on recent theoretical work on the conceptual complexity associated with anonymity (Asenbaum 2018, 459), and within this the distinctive discursive value of pseudonymity (Moore 2018, Veliz 2018). Our second contribution is our focus on the dimension of argumentation and justification, which we analyse through a measurement of the cognitive complexity of comments. We do not claim to address the overall deliberative quality of comments, since deliberative quality is typically taken to encompass dimensions ranging from rationality, interaction or exchange of arguments, equality, civility, constructiveness, and reference to a common good (Friess and Eilders 2015; Stromer-Galley 2007). However, mutual justification is at the core of a wide range of approaches to the study of deliberation,

whether ‘type I’ and ‘type II’ deliberation (Bächtiger et al. 2010). We will discuss this further, after we have set out our approach to identity rules.

Disaggregating Anonymity

Our study sets out not simply to compare commenting behaviour under anonymous and real-name conditions, but also, as we noted above, to thematise pseudonymity. Before describing how we gathered the data, a few words are in order about how we define these terms. Following Fredheim and Moore (2015) and Moore (2018) we disaggregate online identity into three main factors: traceability, durability and connectedness. Traceability means that observers can link online statements and behaviour to real persons. In our context it is important to note that traceability is distinct from the disclosure of identity to fellow commenters. You can make comments anonymously or under a pseudonym and yet it may be possible (with some effort) for advertisers or the security services to trace those comments to the person that made them. Many scholars are concerned about online anonymity in the dimension of traceability (e.g. Zarsky 2004), and seek mechanisms by which online users can remain ‘unreachable’ (Nissenbaum 1999) or ‘untraceable’ (Froomkin 1995) by advertisers or public authorities. Traceability is undoubtedly important to deliberation. It has the potential to chill communication, in so far as it creates a risk of exposure and retaliation from actors who object to certain forms of speech. Yet it also creates the possibility of accountability for abusive or otherwise unlawful behaviour. Depending on whether we have in mind examples of misogynist abusers (Nussbaum 2010) or political dissidents (Zarsky 2004), we might emphasise the value or the dangers of being untraceable. But in either case, traceability is to one side of our concerns in this paper, which turn on the identifiability of commenters *with respect to other commenters in the forum*. Traceability does not help us grasp the relation be-

tween disclosure of identity and discussion quality, so in our analysis we draw primarily on the distinction between durability and connectedness.

Durability refers to the ease or difficulty with which identities can be acquired and changed. A durable identity need not be real-name, but it must be stable over time within a particular context. This dimension has come to the fore in the context of online communication. The concept of durability opens up a distinction between easy anonymity, in which actors are able to easily create new and multiple identities, and stable or durable pseudonyms. It is this stability that grounds the possibility of a limited 'internal' communicative accountability (Moore 2018). As Resnick and Friedman put it, cheap pseudonyms create 'opportunities to misbehave without paying reputational consequences' (2001: 173). Where hurdles such as registration and verification are introduced, it remains possible to create new identities, but it becomes a little harder and more time-consuming. Users are more likely to stick with a particular name, exposing them to the reputational consequences of their behaviour. Where new pseudonyms are easy to create, online identities are disposable; if you acquire a reputation for abusive or untrustworthy behaviour you can just create a new pseudonym and start again. There is some empirical support for this suggestion in findings finds that pre-registration of users can improve the quality of comments (Santana 2014, 28).

Connectedness refers to bridging and linking communication across different social contexts. This dimension can be illustrated by rules or norms of non-disclosure about who said what in a particular deliberative context, designed to block the connection of persons to statements, of which the Chatham House rules are a well known example. In the online context the use of real-names opens the possibility of connectedness in the sense that your statements can po-

tentially become known to your family, colleagues, friends and other social groups with whom you are associated. Connectedness enables statements to be attributed to particular individuals and thereby travel with them into different social contexts. Connectedness works in two directions. It involves the connection of your statements across different social contexts (your comments on a sports team appearing on your professional network). And it involves the exposure of more information about you to interlocutors in a particular context (if commenters on a news site can see your face, name, and perhaps also background, professional position and so on then you are more connected in our terms). Connectedness, finally, admits of degrees. The greater the potential for linking your statements across social contexts and the greater the revelation of personal details to your immediate audience, the more connected you are.

This gives us three modes of identification.

1. Neither durable, nor connected;
2. Durable but not connected; and
3. Durable and connected.

The first two categories are usually described as 'anonymous' both in everyday language and much of the literature reviewed above, and they do indeed share the quality of enabling users to act publicly while concealing aspects of their identity. However, we prefer to describe these modes in terms of durability and connectedness, not least because neither of these modes necessarily involves a complete absence of identifers.[3] There are two further points to note. First, both connectedness and durability admit of degrees. What we are describing below as a real-name environment means that some aspects or cues to your real identity avail-

[3] We thank an anonymous reviewer for pressing this point on us.

able to interlocutors in the forum, and your statements in the forum are attributable to you in other social contexts, that is, that they display higher degrees of connectedness. Second, we recognise that these types are not perfectly recreated in any particular online environment. However, the changes to the commenting architecture made by the Huffington Post over two years from January 2013 enabled us to take a look at something that approximates quite well to these three modes of identification.

Case Study: Changing Identity Rules on the Huffington Post

In our period HuffPo twice changed the identity rules governing its commenting space. The first change took place in December 2013. Prior to December, the platform allowed users to comment under any chosen name, and they could easily comment under different names if they wished. The platform experienced aggressive 'trolling' by users operating multiple or spoof accounts. HuffPo found it was investing too much in policing this increasingly unruly space, and decided to regulate its forum by requiring new users to authenticate their accounts through Facebook.[4] On the face of it, little changed: in the new, durable pseudonymous environment, usernames were still not required to be the user's real name (though some took the option of appearing under their Facebook name and avatar), but behind the scenes Facebook's database helped weed out fake accounts. In June 2014 HuffPo changed to commenting through Facebook, meaning that HuffPo user profiles were replaced by Facebook profiles in a 'real name' environment. In this phase, comments appear below the line of the news article under the user's Facebook name and avatar, as well as — depending on a user's privacy settings — appearing simultaneously on their Facebook page.

---

[4] The new policy was announced here: https://www.huffpost.com/entry/why-is-huffpost-ending-an_b_3817979?guccounter=1, accessed 3 Nov 2019.

The changes made by the Huffington Post allow a series of before and after analyses. We used a script to systematically collect all the conversations below each article. We unravelled each comment thread, recursively checked for, downloaded, and processed any responses to each comment. Having inspected the data, we found that content from the 'featured blog posts' sidebar tended to attract very low commenting levels and therefore introduce a source of noise. For this reason, we excluded texts from the 'featured blog posts' sidebar. In this way we collected more than 45 million comments on more than 50,000 articles featured on the HuffPo front page in the period January 2013 - March 2015.

Method: Measuring Cognitive Complexity

The measurement of deliberative quality has tended to involve qualitative analysis of text or observation of communicative exchanges in light of a range of normative criteria, which typically include practices of justification and argumentation, reciprocity and mutual respect, sincerity, constructive politics, and free participation (Gerhards, 1997; Stromer-Galley, 2007; Steiner, et al., 2004). In our study, however, we focus primarily on the dimension of reason-giving or argumentation. One reason for our restriction is that justification, or asking for and giving reasons, is a common element in an otherwise diverse set of conditions for what makes for good deliberation. We assume that the argumentative dimension is at least necessary, albeit not sufficient, for the deliberative quality of large-scale public discourse (Gastil, 2008).

In order to operationalise (argumentative) deliberative quality, we make use of cognitive complexity (CC) — a psycho-linguistic indicator that measures language complexity (Sued-

feld, et al., 1992; Suedfeld, 2010; Schroder, et al., 1967; Gruenfeld, 1995; Tetlock, 1983; 1984; Abe, 2012; 2011). While it does not directly measure deliberative quality, it serves as a capable proxy. Drawing on empirical comparisons of this method with qualitative analysis of discursive exchanges using tools such as the Discourse Quality Index, CC correlates strongly with the justificatory dimension of deliberation (Beste & Wyss, 2014, Jennstål 2019). We can, therefore, draw on conceptual and empirical validation to warrant our operationalization.

While this approach does not allow us to show the degree to which communicative exchanges exhibited respect or reciprocity — we would not extrapolate from CC to, say, respectful conduct, unless there is a unitary or monolithic conception of deliberation tying together the internal components (e.g. justification and respect) of the discourse ethics—it does permit us to analyse data at large scale. Furthermore, the CC approach is particularly well suited to capture the argumentative dimension of deliberation. We acknowledge that argumentation, defined as the exchange of reasons for or against something (Manin 2005), is perhaps not sufficient for good deliberation; however, it is by all accounts a necessary condition. As a result, recent analyses of political deliberation have started to utilize CC for the purpose of tackling deliberative quality (Brundidge, et al., 2014; Wyss, et al., 2015).

According to the psychological literature, CC refers to the degree of perception multidimensionality of a given problem (van Hiel & Mervielde, 2003; Schroder, et al., 1967; Suedfeld, et al., 1992). It essentially measures how individuals' thought processes are constituted rather than about what actors actually say about a problem, e.g. how logical an argument is. CC consists of two dimensions: (1) differentiation and (2) integration (Owens & Wedeking, 2011). The former refers to the quantity of information, as well as the corresponding problem

dimensions, facets and layers that are considered by a speaker or writer to relate to a given topic, capturing the degree of sophistication with which an individual perceives and conveys an issue. Thus, CC provides cues concerning the breadth of an individual's perspective. Integration, on the other hand, captures whether perceptions of connections, dependencies and other relationships are recognized by a person with respect to a given topic and its substantive sub-dimensions. It refers to the depth of understanding of a speaker or writer, e.g. whether interdependencies that are encapsulated in the problem under discussion (such as externalities, side effects, conflicting values, etc.) are taken into account. Hence, both differentiation and integration tackle disparate dimensions of sophistication: breadth and depth (Schroder, et al., 1967).

When measuring CC, however, differentiation and integration are conceived as an aggregate (Owens & Wedeking, 2011). Originating in the works of 1960s psychologists, the target of the holistic concept (that is the aggregate of differentiation and integration) is to capture an individual's capacity for opinion "flexibility, high levels of information search, and tolerance for ambiguity, uncertainty, and [accepting] lack of closure" (Suedfeld, 2010, p. 1670). High levels of CC are, thus, indicative of an individual's cognitions being embedded, organized and categorized within a dense intellectual system. At the other end of the scale, the CC score diminishes to the extent that the cognitions are narrow, superficial and fragmented. For example, CC has been found to negatively correlate with ideological rigidity, as individuals with a fixed set of beliefs and corresponding preferences are less able to incorporate conflicting new information into their argumentative repertoire (Tetlock, 1984; 1983). Thus, emancipatory ideals such as anti-authoritarianism as well as opinion moderation—both key components of Critical Theory and deliberative politics—are integral parts of high CC scores.

Usually, CC is assessed manually through the labour-intensive reading of documents by trained coders. However, fully automated analyses have become a viable approach to consider digital data in particular (Abe, 2012; 2011; Brundidge, et al., 2014; Wyss, et al., 2015). This is especially true when large quantities of data are analysed and manual coding becomes virtually impossible. As for the means of measuring CC in an automated fashion, the Language Inquiry and Word Count (LIWC) dictionary has been shown to yield valid results in a variety of formal and informal settings (Tausczik & Pennebaker, 2010). Since recent studies have also used the LIWC dictionary to gauge CC in a formal context such as prepared speeches and parliamentary debates, we will use it to analyse political online discussions. After careful validation of manifold operationalisations (Beste & Wyss, 2014), we apply the formula suggested by Owens and Wedeking (2011). They use a combination of complex words (words equal or above six letters in length), causation words (e.g. “because”), as well as expressions indicating personal thought processes (e.g. “believe”), discrepancies between the is and ought status (e.g. “should”), inhibition (e.g. “obstacle”), tentative conclusions (e.g. “perhaps”), inclusivity (“and”), exclusivity (e.g. “except”), negations and certainty (e.g. “absolutely”). We then calculate the Z-standardized values for each unit of analysis by inserting the above-mentioned LIWC indicators into the following formula in order to obtain the CC score:

$$CC = Z(\text{six-letters}) + Z(\text{causation}) + Z(\text{insight}) + Z(\text{discrepancy}) + Z(\text{inhibition}) + Z(\text{tentative}) + Z(\text{inclusive}) - Z(\text{exclusive}) - Z(\text{negations}) - Z(\text{certainty})$$

The unit of analysis are individual user comments, which can be theorized as “speech acts” in the online environment. By applying an automated dictionary-based approach, CC applies a

similar “bag-of-words” (BOW) approach: speech is not parsed in its entire functional differentiation (say, the grammatical structure), but speech acts are rather processed as agglomerations of words. Those words, in turn, have psycho-linguistic properties, which are indicative of the speeches’ complexity. For instance, the word “because” is considered to signal a justification in a speech act, regardless of its embeddedness in the sentence structure.

BOW methods come with specific disadvantages. Since language merely is interpreted as a sum of expressions, there are no contextual considerations. For instance, automatic BOW analyses fail to detect topical or cultural specifics when it comes to, say justification levels. Normally, therefore, it is imperative to adapt the dictionaries to the domain they are applied to. In the case of the LIWC and CC, however, the linguistic indicators in the formula are such that the in- and extension of the concept CC can be maintained and validly measured in a rather universal fashion (Owens & Wedeking, 2011). This is because, firstly, the psycho-linguistic character of the LIWC dictionary in conjunction with the CC concept has already been successfully applied in a host of different political and non-political contexts (Abe 2011, 2012, Tausczik & Pennebaker, 2010). Secondly, the LIWC is less vulnerable for error terms and systematic distortions when comparing cases *within the same domain*. Thirdly, we performed random sample checks of the codings to ascertain the sensibility of the automated approach. Hence, we are confident that we can validly capture Cognitive Complexity for the specific case of Huffington Post user comments.

Results

For the verification of our hypotheses we collected 45m comments posted on the Huffington post website between January 2013 and May 2015. In order to analyse the evolution of cognitive complexity over time, we aggregated the comments of the actively commenting Huffington post users to monthly averages. After aggregation, we end up with 2.7m user-month observations from in total 336.574 users. That is to say that we only considered the observations that ground on at least three user comments. This filter was applied to improve the robustness of the cognitive complexity score.[5] By this procedure, we end up with 2.7m user-month observations from in total 336 574 users, and in total 43 554 480 comments.

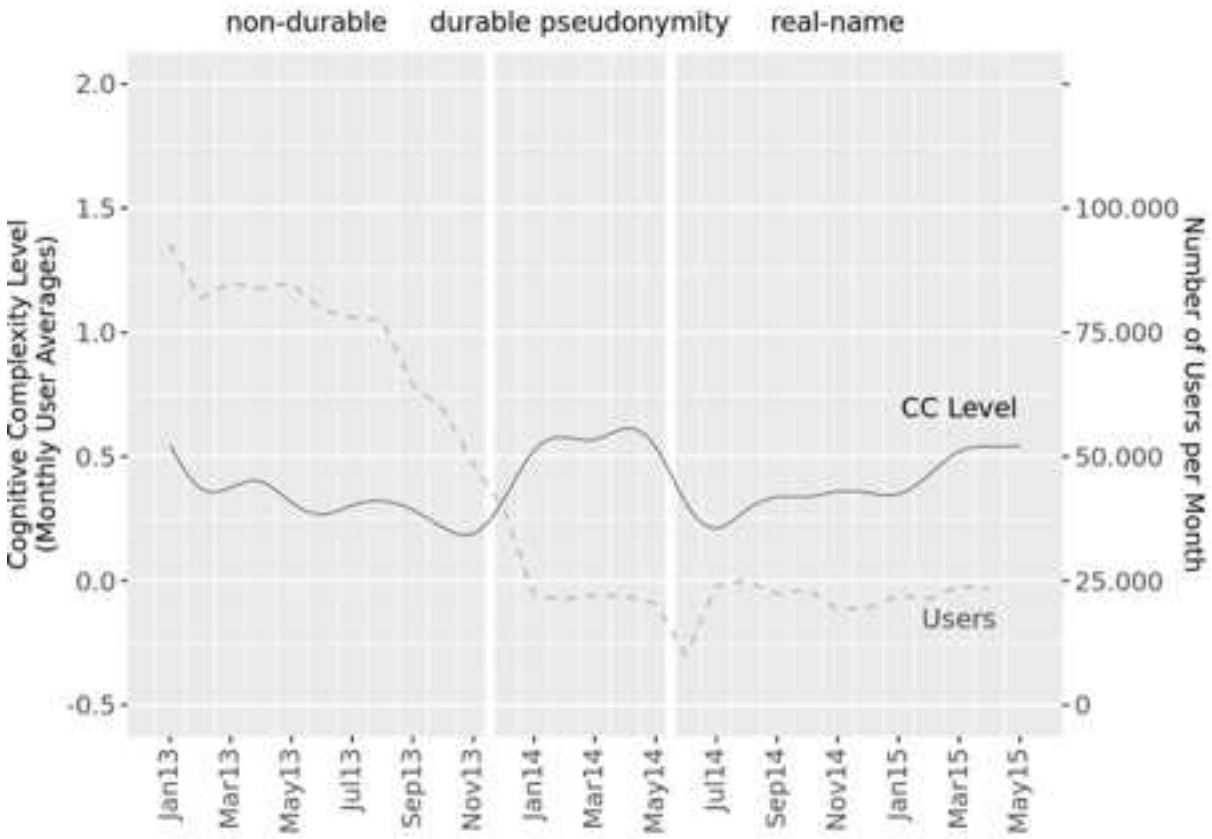
For a first overview over the data, we plot the courses of our cognitive-complexity variable and the number of active users over each observed month (see figure 1). Two observations are particularly striking. During the observed time, we recognize a significant decrease in the number of users posting comments (grey curve). In the years 2014 and 2015, the comments posted on the Huffington Post website were only a fraction of the numbers in 2013. Certainly, this decrease can be seen as a success indicator: The first policy reform in December 2013 helped to remove fake accounts and made it harder for blocked users to reappear under a new account.

The second striking observation is that the cognitive complexity level (black curve) was not

[5] A quarter of the users did not write more than 28 words per month which raises doubts regarding the robustness of the resulting cognitive complexity score. By applying above-mentioned frequency filter, the problem can be mitigated sharply. The filter raises the first word-count quartile from 28 to 147 words. Our subsequent analyses, however, shows that this filter does not influence our findings. The identified effects are astonishingly similar when we apply a more intervening filter (i.e. a 10 comments per month threshold) or when we do not apply any filter at all.

constant over time and seems to be influenced by the policy reforms. Cognitive complexity was highest in the pseudonymity regime in the first half of the year 2014. Under this regime, users required a verified account to post comments on the Huffington Post website, while their public profile remained untouched. This first analyses suggests that the relatively soft reform had an positive effect on the cognitive-complexity index, which increased about 0.6 points. The index decreased again after the second policy change. Hence, under pseudonymity, the average comments found on the huffington post website were written in a more complex manner than the comments found in the surrounding phases.

Figure 1: User-Averages of Cognitive Complexity over Time



The plot shows the number of commenting users (grey dashed line) as well as the level of

cognitive complexity over time (black solid line). The latter line has been estimated by a locally weighted regression model (LOESS; N: 1.2 Mio; k: 20).

For us, an intriguing question is, whether this Cognitive Complexity pattern can be attributed to shifts in individual behavior. Did users indeed write comments in a higher quality? This question cannot be answered by the above analysis, because the reported increase of the cognitive complexity level could also be triggered by a participation bias. That is, a numerical reduction of bots and abusive users could also have resulted in an increase in cognitive complexity. To sort out this alternative effect, we are going to reduce the dataset to users, who posted comments in more than one phase. In the next section, we only consider users who posted under pure anonymity as well as under pseudonymity, which allows us to get a more detailed picture on the effects of the 1st. policy reform. In the subsequent section, we are going to gauge behavior change triggered by the 2nd reform. There, we limit the analyses to users that posted comments in the real-name identify period as well as at least in one of the former periods.

1st Policy Reform: From Non-Durable to Durable Pseudonymity

To gauge whether comment quality changes during the first policy reform, we only consider users that posted comments both before and after the first reform. Following this procedure, we end up with a data set made up of 47.657 unique users.

Instead of just plotting the raw data, we estimate a statistical model that describes the course of measured Cognitive Complexity level. It turns out that a simple but adequate model is the

following OLS model (Formula I):

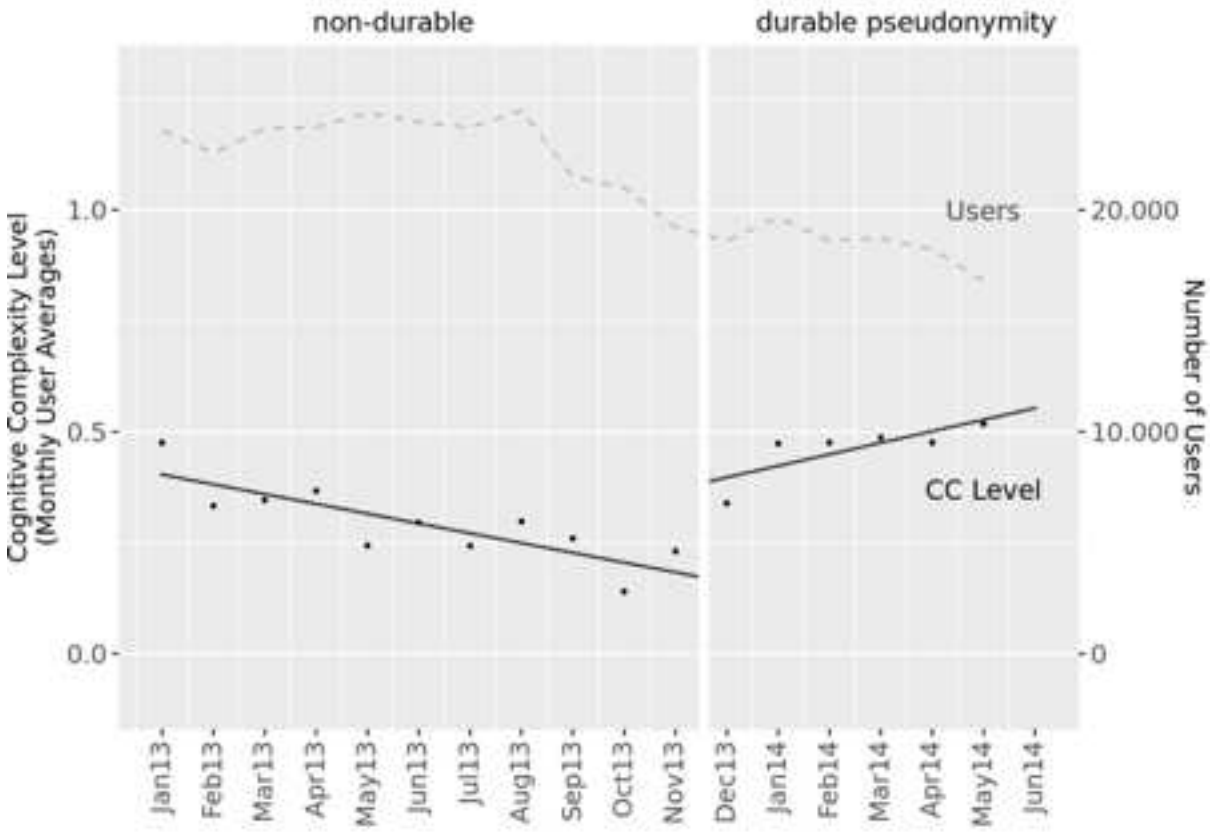
$$\text{Formula I: } CC = \beta_0 + \beta_1 R + \beta_2 M + \beta_3 RM + e$$

In the above model, R stands for reform and represents a dummy variable which is coded as "0" for all observations before, and "1" for all the observations after the first reform. M is a normalized variable measuring the time distance from the data point to the first reform. Finally, the model also features an interaction term between the reform dummy and the distance variable M. This enables to model a potential long-term effect of the reform.

By estimating this model (N: 362.591, F: 243***), we again find that the comment quality measured by Cognitive Complexity increased just after the first policy reform (the coefficient of the reform dummy equals 0.24***). Hence, just after the reform, the posted comments are in average on a higher complexity level than the comments before the reform. The result suggests that the durability of user profiles indeed has a positive effect on comment quality. Because, the measured effect size of 0.24 is about the same strength as it has already been suggested in figure 1, we think that there is not much room for a potential alternative mechanism (i.e. the participation bias) to increase the CC-level. Apparently, regarding the cognitive complexity level, people that posted in the first period do not differ that much from the people posted comments in the second phase. Last but not least, the reform seems to have also a small but significant long-term impact. The interaction term between reform dummy and time distance equals 0.05***. The model is visualized in figure 2, which makes apparent that the

operators of the Huffington Post could achieve by this policy reform a turn-around regarding comment quality.

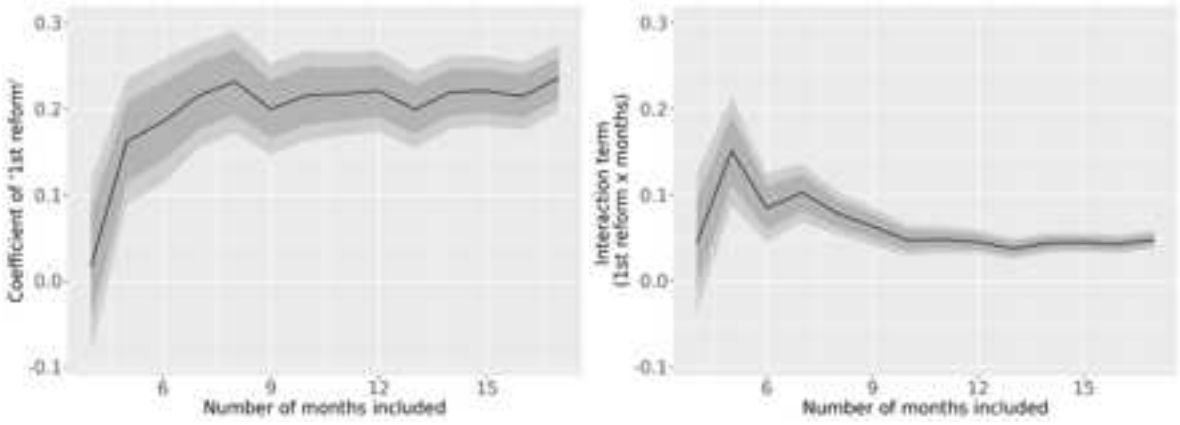
Figure 2: Modelling CC-Level during 1st Reform



The plot shows the number of commenting users (grey dashed line) as well as the level of cognitive complexity over time. The latter is indicated as the monthly mean of user averages (black dots) and the cognitive complexity level predicted by above specified model (black solid line).

How robust is this model? We tested the robustness of our results using different model specifications and a bandwidth test. The latter is a straightforward procedure to gauge model robustness. In above model, we considered the largest possible time range before and after the reform (from January 2013 to May 2014 standing for a bandwidth of 18 months). It is helpful to learn, whether the estimated coefficients change, when we reduce month by month the bandwidth. According to Figure 3, both coefficients remain at a similar level when we shrink the bandwidth. This result supports that the estimated model is stable even when we significantly change the range of observation.

Figure 3: Bandwidth tests: changes of coefficients when number of months included varies



Both, the coefficient of the dummy variable "reform" as well as the coefficient of the interaction term are significantly positive independently of the number of surrounding

months included in the model estimation. The shaded band represents the 95% (dark) and 99% (bright) confidence interval.

2nd Policy Reform: From Durable Pseudonyms to Real-name Identity

We now turn to the second reform and gauge whether the outsourcing of the comment section to Facebook provoked a change in individual behavior. The analysis of this second reform comes with a technical difficulty: Because user profiles were transferred to Facebook, we cannot directly connect the active users of this third phase with the user profiles in the earlier phases. We could establish such a connection only for a subset of users that used already before the second reform the Facebook avatar for their Huffington Post profile. By making use of this link, a total of 8.805 users could be identified as users who posted before and after the second reform.

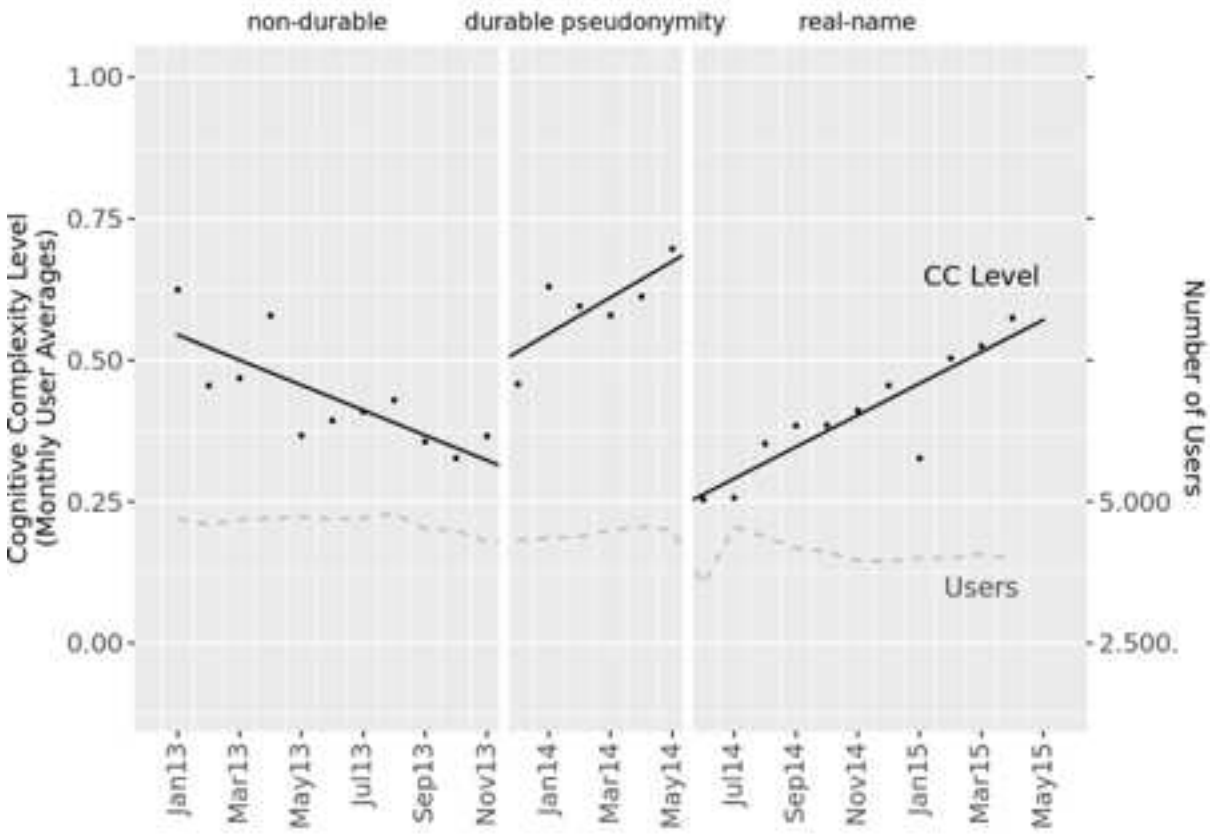
The formula of the model used to analyze this second reform looks slightly different to the one before (Formula II). It features two dummy variables and two interaction terms. Similar as before, $R_{1/2}$ stands for the first or the second reform and represent dummy variables coded as "0" for all observations before, and "1" for all the observations after the specific reform. M is a normalized variable measuring the time distance of the data point to the second reform:

Formula II: $CC = \beta_0 + \beta_1 R_1 + \beta_2 M + \beta_3 R_1 M + \beta_4 R_2 + \beta_5 R_2 M + e$

After estimating this second model (N: 104.514, F: 38.8***), we again find support for the

individual effect of the first reform (the coefficient of the 1st reform dummy equals 0.21***). Most importantly, the coefficient of the second reform is negative. Hence, it seems that the outsourcing to Facebook resulted in a diminishing level of cognitive complexity. Furthermore, both interaction terms are significantly positive, suggesting that in both cases, a positive long-term effect occurred. This second model is visualized in the figure 4, which makes apparent that the quality of comments of the Huffington Post reached a peak during phase 2.

Figure 4: Modelling CC-Level during the Reforms

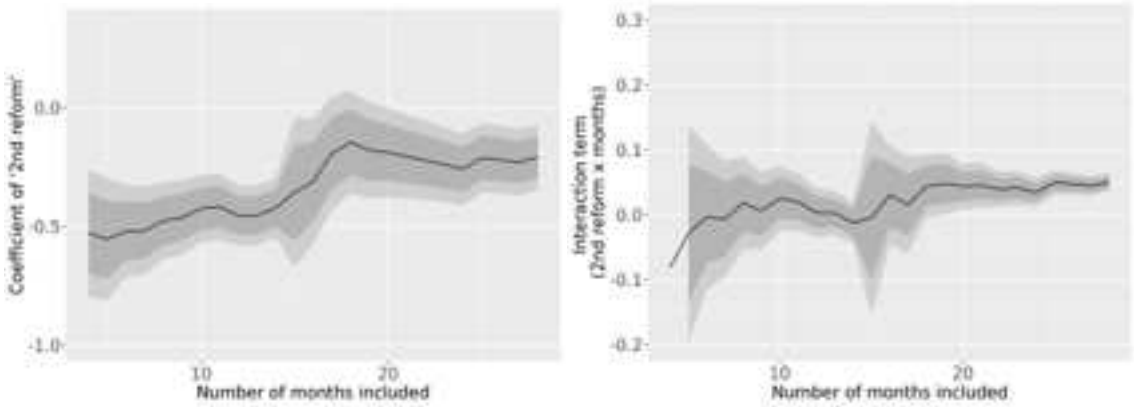


The plot shows the number of commenting users (grey dashed line) as well as the level of cognitive complexity over time. The latter is indicated as the monthly mean of user aver-

ages (black dots) and the cognitive complexity level predicted by above specified model (black solid line).

Again, we test model robustness by using different model specifications, distinct filter criteria, and bandwidth tests. While above model is estimated over all available months (bandwidth of 30), we are interested to learn how coefficients change when we include only a smaller number of months. For estimating the models with bandwidth lower than 16 months (January 2014 until Mai 2015), we removed from the model specification the first interaction term and the first dummy variable. According to Figure 5, the coefficient of the dummy variable remains significantly negative over a large part of the bandwidth range. This is not the case for the interaction term, which becomes insignificant when bandwidth is less than 19 months. According to this test, it is everything but clear that the second reform triggered a long-term effect.

Figure 5: Bandwidth tests: changes of coefficients when number of months included varies



Only the coefficient of the dummy variable "reform" remains significant positive when number of months included shrinks. The coefficient of the interaction term is not significantly difference to zero as soon as the model includes fewer than 20 surrounding

months. The shaded band represents the 95% (dark) and 99% (bright) confidence interval.

What can we say about the effect size? Is a cognitive-complexity change of 0.24 points a small or a large effect? To better assess the effect size, it helps to look for comparable effects found in previous empirical literature. Wyss et al. (2015) found for the context of Swiss Parliament debates a difference in cognitive complexity of 0.32 depending on whether the speaker is a represent of a Government party or a opposition party. In a study of 2014, researchers found that chronic stress was related to linguistic complexity with ß of −0.25 (Saslow et al. 2014). Considerable higher effects on CC are found in case of the opinion complexity of US supreme court justices. The cognitive complexity drops by more than two index points when justices state dissenting opinions. These studies gave a range set out a range of what effect strength we can expect. Regarding a relatively soft institutional change, the effect size we identify here meets expectations.

Discussion

The expectation that poor discursive behaviour in anonymous environments would be improved if uses had their 'real names down' seems not to be clearly or straightforwardly borne out by this case. The most striking finding from our data set is that the cognitive complexity of comments shows a marked improvement in the shift from non-durable to durable pseudonymity. But cognitive complexity reduces again in the 'real-name' phase, when comments are made under the users' Facebook names. We find this pattern repeated when we restrict the analysis to those present through all three phases under study. Here we raise some potential explanations of this pattern and discuss its significance.

The idea that people behave better with their real names down is primarily underpinned by assumptions about communicative accountability. One potential explanation for the difference between the latter two phases can still draw on the mechanism of communicative accountability. However, following Moore (2018) we can distinguish two sorts of communicative accountability: accountability to the audience within the forum itself on the one hand, and a broader accountability for one's speech and actions that is not limited to a particular discursive context on the other. Moore's discussion of the deliberative potentials of pseudonymity suggests that continuity of identity within a particular discursive context is a necessary condition for a minimal form of communicative accountability, which involves the possibility of making and meeting demands for justification within the forum.

What is introduced during the 'real-name' phase, however, is not only durability, but also connectedness. Adding connectedness brings a shift from an audience of only other commenters and readers on the platform to an audience that might also include your other friends, colleagues, and relatives. With the introduction of commenting through Facebook, users not only get cues about the identity of their interlocutors; they also find — depending on their settings — that their comments appear on their Facebook timeline as well as on the HuffPo comment page. Root comments (comments directly on a particular article) can be cross-posted if users choose. Comments on those comments are confined to the respective comment space. For instance, if a friend comments on a HuffPo article and cross-posts to Facebook, and I reply on Facebook, this will not appear below the line on HuffPo. And if I comment below the line on HuffPo it will not show up as a comment on my friend's Facebook timeline. The process is somewhat confusing, and it means that in the Facebook phase the primary

audience of any given commenter becomes uncertain. In the anonymous and pseudonymous phases we can be confident that comments are oriented to the audience of other readers in the comment space. In the real name space, users may orient more to the audience in the comment space, or to their own friendship group on Facebook, or both. This opens up the possibility that the comments people choose to make may be conditioned by the expected response of their friends.

What does this shift in audience entail? One possibility is that a network composed of family, friends, and colleagues is one in which people are less likely to either (i) encounter difference (as one's networks may reflect a degree of prior homogeneity) or (ii) engage actively with those holding different positions, or both. Regarding exposure to diverse opinions, Bakshy et al. (2015), conducting research for Facebook, claim that Facebook users are in fact exposed to a wide range of diverse opinions on the platform. However, there is evidence that Facebook users are less likely to discursively engage and argue with those with different opinions on the platform. Recent research by Rossini (2018), for instance, finds that justified opinion expression is '80% less likely to occur on Facebook than on news sites' (Rossini 2018, 20). She recognises that discourse on Facebook may be more civil, but she — rightly, in our view — separates the question of civility from that of argumentation. With this in mind, we could say that Facebook is perhaps not where you go to engage the other side, but where you go to talk with your friends about the other side. While this is necessarily speculative, we think that the shift in audience from a space in which users engage with diverse others under terms of structured impersonality to a space in which users engage primarily with self-selected networks of friends is a plausible explanation for the observed reduction in argumentative and justificatory modes of discourse.

Limitations

One limitation of our study is that it is not a true experiment, as there is no overlap between the commenting regimes; it is possible that other factors explain the variation in user behaviour. One such factor might be the electoral calendar, though we do not consider this explanation likely. Second, we know that Huffington Post used both manual and algorithmic moderation in all three phases, but we don't know how the policies changed under the different identificatory regimes. It is possible that some of the observed skew towards higher deliberative quality is due to intensive moderation of the politics section during the pseudonymous phase. Future projects may consider working with moderators, and even running experiments on different modes of moderation.

Third, these platforms are not static. Online news platforms and social media experienced rapid changes in their user base and the ways users engage with them during the two-year period of this study. However, this problem may be limited in so far as our study considers a relatively short period of time, and within that period we find clear changes in activity coinciding with policy changes. We must also keep in mind the likely cultural specificity of the ways in which both Facebook and online news commenting are used. Just as the character of these platforms changes over time, so too it is different in different parts of the world. Our study is limited to the American context.

A final limitation to keep in mind is that in constructing online discussion spaces, there are many important design features that are likely to influence commenting behaviour, including, but not limited to: whether comments are ordered chronologically, or curated by editors, or

upvoted by other users, or some other selection method; whether commenting is real-time or asynchronous; and how moderation is performed. We have focused in this study on identity disclosure, but we are aware that it is not the only or even perhaps the most important factor. Moderation practices in particular are hugely important (Grimmelman 2015). However, given the low apparent cost of changes in identifiability and their professed popularity among social media platforms and news providers seeking easy ways to sanitise their comment spaces, it is important to explore the dimensions and effects of anonymity, pseudonymity and real-name environments in more detail.

Conclusion

Our findings complicate the idea that there is a relatively straightforward trade-off between the benefits and drawbacks of anonymous commenting on the one hand, and the benefits and drawbacks of real-name environments on the other. When it is said that commenters ought not to be anonymous, we argue that this does not necessarily mean they must use their real names or connect more fully to their real-world identities; rather, it means that commenting spaces should be structured so as to support the durability of identities, which can at least open the possibility of greater commitment to the commenting space and enhance the potential for learning effects. By emphasising the distinct qualities of durable pseudonymity, our results suggest an important consideration in the design of online discussion spaces.

We conclude with a brief comment on an aspect of the relation of identity rules and deliberation that strikes us as extremely important even though our particular study has not been able to address it. We did not give ‘traceability’ a central place in our analysis because it seemed

that the capacity of commercial entities or governments to trace particular users could apply to users in all of the phases of our study, and especially in the 'durable pseudonym' and 'real-name' phases, in which the identity of users is guarded by platforms. However, we think that this feature of the infrastructure of the public sphere — that platforms increasingly hold the effective power over the concealment and revelation of identity — is of great importance. Thus, while we hope to have contributed to a better understanding of the relation between identity rules and deliberative quality in a particular online forum, there remain a set of larger unanswered questions: What relations of power are enabled and constrained by practices of anonymity? Who has effective control over the conditions of concealment, revelation or construction of identity? These are, we think, crucial further questions if we are to better understand the many dimensions of the relationship between anonymity and public deliberation.

References

Abe J A A (2011) Changes in Alan Greenspan's Language Use Across the Economic Cycle: A Text Analysis of His Testimonies and Speeches. *Journal of Language and Social Psychology* 3(2): 212-223.

Abe J A A (2012) Cognitive–Affective Styles Associated With Position on War. *Journal of Language and Social Psychology* 31(2): 212-222.

Allen D (2010) Anonymous: On Silence and the Public Sphere. In Austin Sarat (ed.), Speech and Silence in American Law. Cambridge: Cambridge University Press, pp. 106-33.

Asenbaum H (2018) Anonymity and Democracy: Absence as Presence in the Public Sphere. *American Political Science Review* 112(3): 1-14.

Bächtiger A, Niemeyer S, Neblo M, Steenbergen MR, and Steiner J (2010) Disentangling diversity in deliberative democracy: Competing theories, their blind spots and complementarities. *Journal of Political Philosophy* 18(1): 32-63.

Bakshy E, Messing S, and Adamic L (2015) Exposure to ideologically diverse news and opinion on Facebook. *Science,* 348(6239): 1130-2.

Beste S (2016) Studying dynamics of content and form in discourses: Evidence from an exploratory study of US legislative discourse on American–Iraqi relations 1995–2012. *Acta Politica* 51(3): 346–371.

Beste S and Wyss D (2014) *A psychological proxy for deliberation? A theoretical and an empirical exploration of the conceptual boundaries of cognitive complexity and deliberative quality.* Glasgow: Paper presented for the ECPR General Conference 2014.

Black L (2011) *The promise and problems of online deliberation*. Charles F. Kettering Foundation, Inc.

Bowker N and Tuffin K (2003) Dicing with deception: People with disabilities' strategies for managing safety and identity online. *Journal of Computer-Mediated Communication* 8(2), 1 January 2003, JCMC827, https://doi.org/10.1111/j.1083-6101.2003.tb00209.x.

Brundidge J A, Scott R, Choi S, and Muddiman A (2014) The "Deliberative Digital Divide:" Opinion Leadership and Integrative Complexity in the U.S. Political Blogosphere. *Political Psychology* 36(6): 741-755.

Cho D and Acquisti A (2013) The More Social Cues, The Less Trolling? An Empirical Study of Online Commenting Behavior. *Unpublished manuscript*, Available from: http://weis2013.econinfosec.org/papers/ChoWEIS2013.pdf (accessed 23 March 2015).

Christopherson K M (2007) The positive and negative implications of anonymity in Internet social interactions: 'On the Internet, nobody knows you're a dog'. *Computers in Human Behavior* 23(6): 3038–3056.

Chun W (2015) The dangers of transparent friends: crossing the public and intimate spheres. In: Danielle Allen and Jennifer Light (eds), From Voice to Influence: Understanding Citizenship in a Digital Age. Chicago: University of Chicago Press, pp.105–28.

Coleman S and G Moss (2012) Under Construction: The Field of Online Deliberation Research. Journal of Information Technology and Politics 9(1): 1-15.

Collins L and Nerlich B (2014) Examining User Comments for Deliberative Democracy: A Corpus-driven Analysis of the Climate Change Debate Online. *Environmental Communication: A Journal of Nature and Culture* 9(2): 189-297.

Davies T and Chandler R (2011) Online Deliberation Design: Choices, Criteria, and Evidence. In: Nabatchi T, Gastil J, Weiksner G and Leighninger M (eds) *Democracy in motion: Evaluating the practice and impact of deliberative civic engagement* (pp. 103–131). Oxford: Oxford University Press.

Foxman A and Wolf C (2013) Viral hate: Containing its spread on the internet. New York: Palgrave Macmillan.

Fredheim R, Moore A and Naughton J (2015) Anonymity and online commenting: the broken windows effect and the end of drive-by commenting. *Proceedings of the ACM Web Science Conference*, Oxford, June 28–July 1. Available at <https://doi.org/10.1145/2786451.2786459>.

Froomkin A M (1995) Anonymity and its enmities. Journal of Online Law, art. 4.

Gapper J (2011) It is right to curtail web anonymity. Financial Times, August 31 2011. https://www.ft.com/content/f3637672-d31e-11e0-9ba8-00144feab49a (accessed 11 Jan 2019).

Gastil J (2008) *Political Communication and Deliberation.* Thousand Oaks: SAGE.

Gerhards J (1997) Diskursive versus liberale Öffentlichkeit. Eine empirische Auseinandersetzung mit Jürgen Habermas. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 49(1): 1-34.

Goodin R E (2008) *Innovating Democracy.* Oxford: Oxford University Press.

Graham T and Wright S (2015) A Tale of Two Stories from "Below the Line": Comment Fields at the Guardian. *International Journal of Press/Politics* 20(3): 317-338.

Grimmelmann J (2015) The virtues of moderation. Yale Journal of Law and Technology 17: 42–109.

Gruenfeld D H (1995) Status, Ideology, and Integrative Complexity on the U.S. Supreme Court: Rethinking the Politics of Political Decision Making. *Journal of Personality and Social Psychology* 68(1): 5-20.

Habermas J (1996) Between Facts and Norms: A Discourse Theory of Law and Democracy. Cambridge: Polity Press.

Hille S and Bakker P (2014) Engaging the Social News User. Journalism Practice 8(5), 563-572.

Janssen D and Kies R (2005) Online Forums and Deliberative Democracy. Acta Politica 40: 317-35.

Jennstål J (2019) Deliberation and Complexity of Thinking. Using the Integrative Complexity Scale to Assess the Deliberative Quality of Minipublics. Swiss Political Science Review.

Kim J-Y (2006) The impact of Internet use patterns on political engagement: A focus on online deliberation and virtual social capital. *Information Polity* 11(1): 35–49.

Kuran T (1997) Private truths, public lies: The social consequences of preference falsification. Cambridge, MA: Harvard University Press.

Lea M, Spears R and de Groot D (2001) Knowing me, knowing you: Anonymity effects on social identity processes within groups. *Personality and Social Psychology Bulletin* 27(5): 526–537.

Leshed G (2009) Silencing the Clatter: Removing Anonymity from a Corporate Online Community. In: T. Davies and S.P. Gangadharan (eds.), *Online deliberation: Design, research, and practice*, Stanford, CA: CSLI Publications, pp. 223–232.

Levmore S (1996) The Anonymity Tool. University of Pennsylvania Law Review 144: 2191-2236.

Manin B (2005) Democratic deliberation: why we should promote debate rather than discussion. Paper delivered at Program in Ethics and Public Affairs Seminar, Princeton University, 13 October.

Manosevitch E and Walker D (2009) Reader comments to online opinion journalism: a space of public deliberation. *International Symposium on Online Journalism*, 1–30. Available from: http://online.journalism.utexas.edu/2009/papers/ManosevitchWalker09.pdf (accessed 23 March 2015).

Moore A (2018) Anonymity, Pseudonymity, and Deliberation: Why Not Everything Should Be Connected. *Journal of Political Philosophy 26*(2): 169-192.

Nissenbaum H (1999) The meaning of anonymity in an information age. *The Information Society* 15(2): 141–144.

Nussbaum M (2010) Objectification and Internet Misogyny. In: Saul Levmore and Martha C. Nussbaum (eds.), The Offensive Internet: Privacy, Speech, and Reputation. Cambridge, MA: Harvard University Press, pp.68-87.

Owens R J and Wedeking J P (2011) Justices and Legal Clarity: Analyzing the Complexity of U.S. Supreme Court Opinions. Law & Society Review 45: 1027–1061.

Polat R K and Pratchett L (2009) E-citizenship: reconstructing the public online. In: Durose C, Greasley S, and Richardson L (eds.), *Changing local governance, changing citizens*. Bristol: The Policy Press, pp. 193–210.

Reagle J (2015) *Reading the comments: Likers, haters, and manipulators at the bottom of the web*. Cambridge, MA: MIT Press.

Reicher S D, Spears R, and Postmes T (1995) A social identity model of deindividuation phenomena. *European review of social psychology* 6(1): 161–198.

Resnick P and Friedman E J (2001) The social cost of cheap pseudonyms. *Journal of Economics & Management Strategy* 10(2): 173–199.

Rhee J W and Kim E (2009) Deliberation on the net: Lessons from a field experiment. In: T. Davies and S. P. Gangadharan SP (eds), *Online deliberation: Design, research, and practice*. Stanford, CA: CSLI Publications, pp. 223–232.

Rossini, P. (2018). Does it Matter if it's Uncivil? Conceptualizing Uncivil and Intolerant Discourse in Online Political Talk. Paper presented at the 68th Annual Conference of the International Communication Association (ICA), Prague, May 24-28, 2018.

Rowe I (2014) Civility 2.0: A Comparative Analysis of Incivility in Online Political Discussion. *Information, Communication & Society* 18(2): 121-138.

Saslow L R, McCoy S, van der Löwe I, Cosley B, Vartan A, Oveis C, Keltner D, Moskowitz J T, and Epel E S (2014). Speaking under pressure: low linguistic complexity is linked to high physiological and emotional stress reactivity. Psychophysiology 51: 257–66.

Santana A D (2014) Virtuous or Vitriolic?: The Effect of Anonymity on Civility in Online Newspaper Reader Comment Boards. *Journalism Practice* 8(1): 18-33.

Schroder H M, Driver M J, and Streufert S (1967) *Human Information Processing. Individuals and Groups Functioning in Complex Social Situations.* New York: Holt, Rinehart and Wiston.

Sia C, Tan B C Y, and Wei K (2002) Group polarization and computer-mediated communications: effects of communication cues, social presence, and anonymity. *Information Systems Research* 13(1): 70–90.

Steiner J, Bächtiger A, Spörndli M, and Steenbergen MR (2004) *Deliberative politics in action: Analyzing parliamentary discourse.* Cambridge: Cambridge University Press.

Stromer-Galley J (2007) Measuring deliberation's content: A coding scheme. *Journal of public deliberation* 3(1): 12.

Suedfeld P (2010) The Cognitive Processing of Politics and Politicians: Archival Studies of Conceptual and Integrative Complexity. *Journal of Personality* 78(6): 1669-1702.

Suedfeld P, Tetlock P E, and Streufert S (1992) Conceptual/integrative complexity. In: C. P. Smith, ed. *Motivation and Personality: Handbook of Thematic Content Analysis.* Cambridge: Cambridge University Press, pp. 393-400.

Suler J (2004) The online disinhibition effect. *Cyberpsychology & behavior* 7(3): 321–326.

Tausczik Y R and Pennebaker J W (2010) The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* 29(1): pp. 24-54.

Tetlock P E (1983) Cognitive style and political ideology. *Journal of Personality and Social Psychology* 45: 118-126.

Tetlock P E (1984) Cognitive Style and Political Belief Systems in the British House of Commons. *Journal of Personality and Social Psychology* 46(2): 365-375.

Towne E and Herbsleb J (2012) Design Consideration of Online Deliberation Systems. Journal of Information Technology and Politics 9(1): 97-115.

van Hiel A, and Mervielde I (2003) The Measurement of Cognitive Complexity and Its Relationship With Political Extremism. *Political Psychology* 24(4): 781-801.

Véliz C (2018) Online Masquerade: Redesigning the Internet for Free Speech Through the Use of Pseudonyms. *Journal of Applied Philosophy* https://doi.org/10.1111/japp.12342.

Weber P (2014) Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments. *New Media & Society* 16(6): 941–957.

Wyss D, Beste S, and Bächtiger A (2015) A Decline in the Quality of Debate? the Evolution of Cognitive Complexity in Swiss Parliamentary Debates on Immigration (1968-2014). Swiss Political Science Review 21: 636–653.

Zarsky T (2004) Thinking outside the box: considering transparency, anonymity, and pseudonymity as overall solutions to the problems of information privacy in the Internet society. University of Miami Law Review 58: 991–1044.